AD_____

Award Number: DAMD17-00-1-0031

TITLE: Functional Analysis of C-CAM1 Tumor Suppressor Gene by Targeted Gene Deletions

PRINCIPAL INVESTIGATOR: Sue-Hwa Lin, Ph.D.

CONTRACTING ORGANIZATION: University of Texas
                          M.D. Anderson Cancer Center
                          Houston, Texas 77030

REPORT DATE: January, 2001

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
               Fort Detrick, Maryland 21702-5012

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

1

20010620 195

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE January, 2001	3. REPORT TYPE AND DATES COVERED Annual (1 Jan 00 - 31 Dec 00)

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Functional Analysis of C-CAM1 Tumor Suppressor Gene by Targeted Gene Deletions	DAMD17-00-1-0031

6. AUTHOR(S)
Sue-Hwa Lin, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
University of Texas M.D. Anderson Cancer Center 1515 Holcombe Blvd. Houston, Texas 77030 E-MAIL: slin@notes.mdacc.tmc.edu	

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING / MONITORING AGENCY REPORT NUMBER
U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012	

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited	12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

C-CAM is a cell adhesion molecule of the immunoglobulin supergene family. We have recently shown that C-CAM plays critical roles in prostate cancer initiation and progression and that loss of C-CAM is an early event in the development of prostate cancer. Although tumorigenesis studies in mouse xenograft model have suggested the involvement of C-CAM in epithelial cell growth and differentiation, the functional roles of C-CAM in normal prostate development, prostate homeostasis, and prostate tumorigenesis remain unclear. We propose to determine the roles of C-CAM's growth suppressive activity in prostate growth and tumorigenesis by using gene targeting and embryonic stem cell technologies to generate C-CAM knockout mice. We have isolated and sequenced two closely related Ceacam (C-CAM) genes, i.e. Ceacam1 and Ceacam2, from a mouse 129 Sv/Ev library. We have also examined the tissue-specific and embryonic expressions of these mouse Ceacam1 and Ceacam2 genes. Results from this study allow us to design a gene targeting strategy that is specific to Ceacam1 gene and also allow us to perform both straight and conditioned knockout of Ceacam1 gene in parallel.

14. SUBJECT TERMS Prostate Cancer			15. NUMBER OF PAGES 28
			16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

(4)   INTRODUCTION

C-CAM is a cell adhesion molecule of the immunoglobulin supergene family (1). We have recently shown that C-CAM plays critical roles in prostate cancer initiation and progression and that loss of C-CAM is an early event in the development of prostate cancer (2). Although tumorigenesis studies in mouse xenograft model have suggested the involvement of C-CAM in epithelial cell growth and differentiation, the functional roles of C-CAM in normal prostate development, prostate homeostasis, and prostate tumorigenesis remain unclear. Towards the aim of determining the roles of C-CAM's growth suppressive activity in prostate growth and tumorigenesis, we propose to use gene targeting and embryonic stem cell technologies to generate C-CAM knockout mice. Specifically, we plan (1) to determine the roles of C-CAM's growth suppressive function in vivo by generating mice with a targeted deletion of the C-CAM cytoplasmic domain; (2) to determine the roles of C-CAM's growth suppressive function in prostate development and tumorigenesis by generating mice with a prostate-specific knockout of the C-CAM cytoplasmic domain. The proposed work was divided into two Tasks to be carried out in parallel.

Task 1. Generate mice with targeted deletion of C-CAM cytoplasmic domain to determine the roles of C-CAM's growth suppressive function in vivo (months 1-30)

Task 2. Prostate-specific loss of function of C-CAM gene in prostate (months 7-36)

Genetic manipulation of mouse genes in vivo is a powerful approach for understanding the function of a gene, both during embryonic development and in adult tissues. This method requires full knowledge of the genomic structure of the gene of interest. Unlike humans and rats, which each have one C-CAM1 (renamed as Ceacam1 (3)) gene, two Ceacam1-like genes, Ceacam1 and Ceacam2, were identified in mice. This poses potential problems in any attempt to manipulate these genes. First, if the two genes have the same function, deletion of one gene may not produce any phenotypic alteration. Second, if there is no significant difference in the genes' sequences, any genetic manipulation may not be specific for a single gene, or it might be difficult to ascertain which gene was altered. It is, therefore, essential to determine the

4

complete sequences and expression profiles of these related Ceacam genes in the 129sv mouse strain before genetic manipulation performed. We, therefore, isolated and sequenced the Ceacam1 and Ceacam2 genes from a mouse 129 Sv/Ev library. In addition, we also examined the tissue-specific and embryonic expressions of these mouse Ceacam1 and Ceacam2 genes.

(5) BODY (Progress report)

5.1. Studies performed under Task 1--Generate mice with targeted deletion of C-CAM cytoplasmic domain to determine the roles of C-CAM's growth suppressive function in vivo (months 1-30)

5.1.1. Determine the genomic structure of *Ceacam1* and *Ceacam2* genes in 129Sv/Ev mouse:

A 129 Sv/Ev mouse genomic library in λLEX (provided by Dr. Li-Yuan Yu-Lee, Baylor College of Medicine, Houston, TX, U.S.A.) was screened with the 1.6-kb full-length mouse *Ceacam1* cDNA (4) which had been labeled by using the Klenow fragment of DNA polymerase I, random hexanucleotide primers and [α-^{32}P] dCTP (5). Twenty-four positive clones were identified in the initial screening of this genomic library. PCR with Oligo 127 (5'-GTGTCACTCTAGGCTACAGGAAAT-3') and Oligo 1 2 2 (5'-GAGGCCAGCTCCATGTCTCTGCTG-3'), which are specific to the 5' region of mouse *Ceacam1* (6), showed that seven of the clones had the N-terminal *Ceacam1* sequence. Similarly, PCR with Oligo 120 (5'-GAAGTCTGGCGGATCTGGCTCCTT-3') and Oligo 131 (5'-TTGAAGTTCAGGACAGTGTATGCG-3') showed that seven clones had the 3' *Ceacam1* sequence. These clones were isolated by secondary and tertiary screening and characterized by restriction mapping. The nucleotide sequences of the exons and introns were determined by primer walking using specific oligonucleotide primers. Sequencing was performed by the DNA Sequencing Core Facilities at M. D. Anderson Cancer Center with an automated fluorescent DNA sequencer (Applied Biosystems Inc., Ramsey, NJ, U.S.A.).

The seven clones with the 5' of *Ceacam1* had different restriction maps, as did another seven clones with the 3' end. The different restriction maps seemed to reflect different genes. Grouping these clones according to their restriction maps and DNA sequences revealed two

5

distinct sequences, *Ceacam1* and *Ceacam2*. The overlapping genomic clones that spanned *Ceacam1* and *Ceacam2* are shown in Figure 1. DNA sequence analysis of both strands of these two genes revealed that *Ceacam1* and *Ceacam2* contained nine exons each and were approximately 18.3 and 24.65 kb, respectively (Figure 1).

The sizes of the exons and introns and the intron/exon boundary sequences for *Ceacam1* and *Ceacam2* are shown in Table 1 and 2. The first exon, which codes for the first two-thirds of the signal sequence, was 304 bp in both *Ceacam1* and *Ceacam2*. In contrast, exons 2-5 were each about 300 bp, and each coded for one Ig-like domain in both *Ceacam1* and *Ceacam2*. The transmembrane domain was encoded by exon 6 and the cytoplasmic domain by exons 7-9. Consistent with the RNA splicing rule (7), each intron started with GT at the 5' end and ended with AG at the 3' end (Table 1 and 2).

5.1.2. Comparison of *Ceacam1* and *Ceacam2* genes

The overall similarity between *Ceacam1* and *Ceacam2*, including all exons and introns, was about 79.6%. The similarities between the exons and introns of *Ceacam1* and *Ceacam2* are shown in Table 3. The major differences between these two genes were in exon 2, intron 2, intron 5, exon 6, and intron 7, which were 76.9%, 29.7%, 59.8%, 78.5%, and 64.5% similar, respectively. The nucleotide and the deduced amino acid sequences of exon 2 for *Ceacam1* and *Ceacam2* are shown in Figure 2. As shown, most of the sequence substitutions in exons 2, which coded for the first Ig domain, did result in amino acid substitutions in *Ceacam1* and *Ceacam2*. Thus, the amino acid identity between the first Ig domains of CEACAM1 and CEACAM2 was 57.1%, which was significantly lower than the similarity of other domains. In addition, insertions of 418 and 5849 bp in intron 2 and of 1384 and 197 bp in introns 5 and 7, respectively, were found in *Ceacam2* compared with *Ceacam1* (Table 3 and Figure 1C).

5.1.3. Expression of *Ceacam1* and *Ceacam2* messages in mouse tissues

To determine the tissue-specific distribution of *Ceacam1* and *Ceacam2*, we performed semiquantitative RT-PCR on a panel of mouse cDNAs from various tissues by using *Ceacam1* and *Ceacam2*-specific oligonucleotide pairs. A mouse cDNA panel containing first-strand cDNA

6

prepared from mouse tissues and normalized for β-actin expression was purchased from OriGene Technologies Inc. (Rockville, MD, U.S.A.) and used to analyze the expression of *Ceacam1* and *Ceacam2* in various mouse tissues. Oligonucleotides specific to exon 2 of *Ceacam1* (Oligo 116, 5'-AATCTGCCCCTGGCGCTTGGAGCC-3', and Oligo 179, 5'-AAATCGCACAGTCGCCTGAGTACG-3') and to exon 2 of *Ceacam2* (Oligo 117, 5'-AATATGATGAAGGGAGTCTTGGCC-3', and Oligo 180, 5'-AAATTGTCCAGTCAGGACCCTACG-3') were used as primers to detect specific messages for *Ceacam1* and *Ceacam2* by PCR. PCR cycling conditions were as follows: (1) pre-denaturation, 94°C, 3 min for one cycle, (2) denaturation, 94°C, 30 s, annealing, 60°C, 30 s, and extension, 72°C, 2 min for 35 cycles and (3) final extension, 72°C, 5 min. The predicted size of the PCR products was 246 bp. These PCR products were analyzed by agarose gel electrophoresis, transferred onto a nylon membrane and hybridized with ^{32}P-labeled oligonucleotide probes specific to *Ceacam1* (Oligo 181, 5'-AACACTACGGCTATAGACAAA-3') and *Ceacam2* (Oligo 182, 5'-TCTACTACGTCTACAAATGCT-3'), the least similar exon (Figure 2A). The specificities of these two oligonucleotides were confirmed by Southern blot analysis. As shown in Figure 3A, Oligo 181 hybridized to DNA from λ3 (*Ceacam1*) but not from λ13 (*Ceacam2*), whereas Oligo 182 only hybridized to λ13 DNA. Both *Ceacam1* and *Ceacam2* mRNAs were detected, although not in the same tissues. This observation supports the notion that both *Ceacam1* and *Ceacam2* are expressed in mouse (Figure 3B and 3C). High levels of *Ceacam1* message were detected in liver, small intestine, prostate, and spleen (Figure 3B), similar to the tissue expression pattern in humans (8) and rats (9,10). *Ceacam1* message was also detected in heart, kidney, stomach, muscle, skin, and uterus. In contrast, *Ceacam2* messages were detected only in kidney, testis, and, to a lesser extent, spleen (Figure 3C). These observations demonstrated that the *Ceacam1* and *Ceacam2* genes are differentially expressed in mouse tissues.

5.1.4. *Ceacam2* cDNA

Although they were detected in testis, kidney and spleen, *Ceacam2* messages may not contain open reading frames for protein translation. As testis expressed only *Ceacam2*, the cDNAs coding for *Ceacam2* were obtained from testis RNA by RT-PCR. RNA was prepared

from mouse testis by using RNAzol B (TEL-TEST Inc. Friendswood, TX, U.S.A.). The cDNAs coding for *Ceacam2* were obtained from the testis RNA by reverse transcription (RT)-PCR with Oligo 630 (5'-GAATTCAAGCTTAAGAAGCTAGCAGGCAGCAGAGAC-3'), which contains nt −36 to +1 of exon 1, and Oligo 631 (5'-GCGGCCGCCTAATGATGATGATGATGATGATGCTTCTTTTTTACTTCTGAATAAAC-3'), which is complementary to the end of the coding sequence in exon 9 plus seven histidine codons. RT-PCR was performed with Oligo 630 and Oligo 631 according to the procedures provided by the manufacturer (Amersham/Pharmacia Corp., Arlington Heights, IL, U.S.A.). The 1-kb PCR product obtained only hybridized to *Ceacam2*-specific Oligo 182, not to *Ceacam1*-specific Oligo 181, suggesting that the fragment codes for *Ceacam2*. This result is consistent with the tissue distribution of *Ceacam2* (Figure 3C). The PCR product was subcloned into pCRII-topo (Invitrogen, San Diego, CA, U.S.A.) and its sequence was determined (Figure 4). The mouse testis cDNA contained exon 1, 2, 5, 6, 8, and 9, and did have an open reading frame of 273 amino acids (Figure 4B).

5.1.5. Expression of *Ceacam1* and *Ceacam2* during mouse embryonic development

The expression of *Ceacam1* and *Ceacam2* was examined in mouse embryos at 8.5, 9.5, 12.5 and 19 days of embryonic development by RT-PCR hybridization as described above. Using *Ceacam1*-specific Oligo 181, we observed age-related differences in mRNA levels: we detected the message at day 8.5, it disappeared at day 9.5-12.5, and it reappeared at day 19 (Figure 5A). In contrast, no hybridization signal was detected when *Ceacam2*-specific Oligo 182 was used. This result suggests that the expression of *Ceacam1*, but not *Ceacam2*, is developmentally regulated.

5.2. Studies performed under Task 2--Prostate-specific loss of function of C-CAM gene in prostate

5.2.1. Development of gene targeting strategy

Because targeted gene deletion is performed in the 129sv mouse, we determined the complete genomic structure and DNA sequences of mouse *Ceacam1* and *Ceacam2* genes in a

129Sv/Ev mouse genomic library. We also examined the expression patterns of these two genes in adult mouse tissues and during embryonic development. Several conclusions can be made from this study. First, the *Ceacam1* and *Ceacam2* genes contained sufficient sequence differences that targeted gene deletion specific to either gene should be feasible. Second, functional redundancy may not be a problem when only one of these two genes is deleted because although *Ceacam1* and *Ceacam2* were highly homologous, they had different tissue expression patterns. Third, *Ceacam1* is probably more important than *Ceacam2* in the mouse because *Ceacam1's* tissue expression pattern was similar to those of the single Ceacam genes in humans and rats. Conservation of the expression profile among these different species suggests that *Ceacam1's* function may be essential. Fourth, *Ceacam2* plays no role in embryonic development; it is not expressed in the mouse embryo.

The existence of the highly homologous second *Ceacam* gene, *Ceacam2*, in mice raises the possibility that *Ceacam2* is a pseudogene. However, our results indicate that *Ceacam2* is probably not a pseudogene because *Ceacam2* contained a complete set of exons and introns typical of a *Ceacam* gene; *Ceacam2* was transcribed in mouse as evidenced by the presence of *Ceacam2* message in several tissues; and *Ceacam2* message contained an open reading frame of 273 amino acids in testis, as it has been shown in the CMT-93 mouse rectal carcinoma cells (11). These findings raise the interesting question of *Ceacam2*'s function. One of the functions of CEACAM1 is inhibition of tumor growth (12,13). This suggests that this protein may play an important role in regulating cell growth and differentiation. Structural and functional analyses of rat CEACAM1 revealed that the tumor suppressive function requires a long cytoplasmic domain generated by alternative splicing (13-15). Because the cytoplasmic domains of CEACAM1 and CEACAM2 were identical, CEACAM2 may also have growth-suppressive activity. However, the role of CEACAM2 in testis, which is composed of rapidly dividing cells in spermatogenesis, is not clear. Thus, the *in vivo* function of CEACAM2 may not be revealed until its gene is deleted in the mouse.

Knocking out a gene can have no phenotypic effect if related genes have similar functions. We showed that *Ceacam1* and *Ceacam2* were expressed in different tissues in the

mouse. *Ceacam1* message was detected in tissues rich in epithelial cells which is consistent with expression of the CEACAM1 homologue in rats and humans (16-18). In contrast, *Ceacam2* message was abundantly expressed in testis, which does not express *Ceacam1* message. In addition, *Ceacam2* message was undetectable in the mouse embryo, whereas *Ceacam1* message was developmentally regulated. Thus, it appears that the two mouse *Ceacam* genes are not functionally redundant. However, we cannot rule out the possibility that the loss of *Ceacam1* expression may upregulate expression of *Ceacam2*. These studies would have to await targeted gene deletion of these two related genes.

The results of this study provide important information for designing gene targeting strategies for functional studies of *Ceacam* genes, by either deleting or introducing mutations in the *Ceacam* genes in the mouse germline. The differences in the two DNA sequences can be used to target a specific gene. As *Ceacam1* lacked the 418 and 5849 bp insertions in *Ceacam2* intron 2 and the 1384 bp insertion present in *Ceacam2* intron 5 (Figure 1C), gene-targeting vectors containing introns 2 or 5 of *Ceacam1* or *Ceacam2* could be used to achieve selective homologous recombination in the desired gene. If mutant mice with only one *Ceacam* gene deleted survive, then they can be further crossed to generate mice deficient in both *Ceacam* genes. However, as *Ceacam1* and *Ceacam2* are both on chromosome 7, it is possible that they are too close to construct double gene knockout mice by crossing the single gene knockout mice.

5.2.2. Gene targeting strategy

Based on the information from the genomic characterization of Ceacam1 and Ceacam2 gene, we decided to delete the cytoplasmic domain of Ceacam1. This strategy is due to the following reasons: (1) The Ceacam1 gene is functionally more important than Ceacam2 gene; (2) The entire Ceacam1 gene, which is around 20 kb, is too large to delete; (3) The cytoplasmic domain of CEACAM1 is critical for tumor suppressor function. The knockout construct planned for is a conditional knockout construct deleting exon 7~9. This way we will have the option of deleting this region both in ES cells and in the mice. For straight knockout, which is the goal of Task 1, we will transfect the cre recombinase expression plasmid to the ES cells and screen for

10

ES cells that have deletions in exon 7-9 of its Ceacam1 gene before injection the ES cells to blastocysts. For the conditional knockout, which is the goal of Task 2, the ES cells that has targeting vector will be injected to blastocysts and the Ceacam1 gene will be deleted in prostate by crossing with transgenic mouse carrying probasin-driven cre recombinase. Therefore, Task 1 and Task 2 are being performed in parallel.

(6) KEY RESEARCH ACCOMPLISHMENTS

- ❖ Isolate and complete sequence two closely related Ceacam genes, i.e. Ceacam1 and Ceacam2, from a mouse 129 Sv/Ev library.
- ❖ Examine the tissue-specific and embryonic expression of these mouse Ceacam1 and Ceacam2 genes.
- ❖ Design knockout strategy that allows both Task1 and Task2 being carried out in parallel.

(7) REPORTABLE OUTCOMES

1 manuscript accepted for publication in Biochem J.: Han, E., Phan, D., Lo, P., Poy, M. N., Behringer, R., Najjar, S., and Lin, S.-H.: Difference in tissue-specific and embryonic expression of mouse Ceacam1 and Ceacam2 genes.

(8) CONCLUSION

We propose to determine the roles of CEACAM's growth suppressive activity in prostate growth and tumorigenesis, we plan to use gene targeting and embryonic stem cell technologies to generate CEACAM knockout mice. We have isolated and sequenced two closely related Ceacam genes, i.e. Ceacam1 and Ceacam2, from a mouse 129 Sv/Ev library. We have also examined the tissue-specific and embryonic expressions of these mouse Ceacam1 and Ceacam2 genes. Results from this study allow us to design a gene targeting strategy that is specific to Ceacam1 gene and also allow us to perform both straight and conditioned knockout of Ceacam1 gene in parallel.

(9) REFERENCES

1.   Beauchemin, N., and Lin, S.-H. (1998) *Role of C-CAM as a tumor suppressor.* Cell adhesion and communication mediated by the CEA family: basic and clinical perspectives. (Stanners, C., Ed.), Harwood Academic Publishers, Amsterdam

2.   Lin, S.-H., and Pu, Y.-S. (1999) *Seminars in Oncology* **26,** 227-233

3.   Beauchemin, N., Draber, P., Dveksler, G., Gold, P., Gray-Owen, S., Grunert, F., Hammarstrom, S., Holmes, K. V., Karlsson, A., Kuroki, M., Lin, S.-H., Lucka, L., Najjar, S. M., Neumaier, M., Obrink, B., Shively, J. E., Skubitz, K. M., Stanners, C. P., Thomas, P., Thompson, J. A., Virji, M., von Kleist, S., Wagener, C., Watts, S., and Zimmermann, W. (1999) *Exp. Cell Res.* **252,** 243-249

4.   Williams, R. K., Jiang, G-S., and Holmes, K. V. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 5533-5536

5.   Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.*, second Ed. (Nolan, C., Ed.), Cold Spring Harbor Laboratory Press, New York

6.   Nedellec, P., Turbide, C., and Beauchemin, N. (1995) *Eur. J. Biochem.* **231**(1)**,** 104-14

7.   Senapathy, P., Shapiro, M. B., and Harris, N. L. (1990) *Methods Enzymol.* **183,** 252-278

8.   Prall, F., Nollau, P., Neumaier, M., Haubeck, H. D., Drzeniek, Z., Helmchen, U., Loning, T., and Wagener, C. (1996) *Journal of Histochemistry & Cytochemistry* **44**(1)**,** 35-41

9.   Odin, P., Asplund, M., Busch, C., and Obrink, B. (1988) *J. Histochem. Cytochem.* **36,** 729-739

10.   Cheung, P. H., Culic, O., Qiu, Y., Earley, K., Thompson, N., Hixson, D. C., and Lin, S. H. (1993) *Biochem. J.* **295,** 427-435

11.   Nedellec, P., Dveksler, G. S., Daniels, E., Turbide, C., Chow, B., Basile, A. A., Holmes, K. V., and Beauchemin, N. (1994) *J. of Virol.* **68**(7)**,** 4525-37

12.   Hsieh, J. T., Luo, W., Song, W., Wang, Y., Kleinerman, D. I., Van, N. T., and Lin, S. H. (1995) *Cancer Res.* **55,** 190-197

13.   Kunath, T., Ordonez-Garcia, C., Turbide, C., and Beauchemin, N. (1995) *Oncogene* **11,** 2375-82

14.   Luo, W., Wood, C., Earley, K., Hung, M.-C., and Lin, S.-H. (1997) *Oncogene* **14,** 1697-1704

15.   Estrera, V. T., Luo, W., Phan, D., Hixson, D., and Lin, S.-H. (1999) *Biochem. Biophys. Res. Commun.* **263,** 797-803

16.   Cheung, P. H., Thompson, N. L., Earley, K., Culic, O., Hixson, D., and Lin, S. H. (1993) *J. Biol. Chem.* **268,** 6139-46

17.   Thompson, J. A., Grunert, F., and Zimmermann, W. (1991) *Journal of Clinical Laboratory Analysis* **5,** 344-366

18.   Hinoda, Y., Neumaier, M., Hefta, S. A., Drzeniek, Z., Wagener, C., Shively, L., Hefta, L. J. F., Shively, J. E., and Paxton, R. (1988) *Proc. Natl. Acad. Sci.* **85,** 6959-6963

Table 1 Intron/exon boundaries of *Ceacam1*

Exon	Size (bp)	Exon 3'		Intron	Size (bp)		Exon 5'
Exon 1	304	ACA G Thr A	gtaaggagatattcc…	intron 1	767	…tctcttccctcttag	CC TCA la Ser
Exon 2	360	CAC C His P	gtaagtaattatctg…	intron 2	1540	…ctattatctgcacag	CC ATA ro Ile
Exon 3	285	ATC T Ile T	gtgagtaacttcttc…	intron 3	467	…ttctgtttgctccag	AT GGT yr Gly
Exon 4	255	CTT G Leu G	gtaagtggatctctg…	intron 4	1727	…tctgtttgtccacag	AG CCA ln Pro
Exon 5	276	ATA T Ile P	gtgagtgacctcgcc…	intron 5	5358	…ttctttccctgacag	TT GAC he Asp
Exon 6	121	GGC GG Gly Gl	gtaggacagtctttc…	intron 6	796	…ctcatatttatttag	G GGA y Gly
Exon 7	53	CAC A His A	gtaagtaaagccaat…	intron 7	587	…ttcttcctcccctag	AT CTG sn Leu
Exon 8	32	AAC AAG Asn Lys	gtgagcactgccact…	intron 8	948	…ctctcatcctttcag	GTG GAT Val Asp
Exon 9	1180						

Table 2 Intron/exon boundaries of *Ceacam2*

Exon	Size (bp)	Exon 3'		Intron	Size (bp)		Exon 5'
Exon 1	304	ACA G Thr A	gtaaggagatattcc…	intron 1	766	…tctcttccctcttag	CC TCA la Ser
Exon 2	360	CAC A His T	gtaagtaattctctg…	intron 2	7807	…acacagtgcagacag	CC CTA hr Leu
Exon 3	285	ATC T Ile T	gtgagtaacttcttt…	intron 3	465	…tcttgtttgctccag	AT GGT yr Gly
Exon 4	255	CTT G Leu G	gtaagtggatctctg…	intron 4	1717	…tctgtttgtccacag	AG CCA lu Pro
Exon 5	276	ATA T Ile P	gtgagtgaccttgcc…	intron 5	6737	…ttctttccctgacag	TT GAC he Asp
Exon 6	118	CGC TG Arg Tr	gtaggacagtctttc…	intron 6	798	…ctcatatttatttag	G GGA p Gly
Exon 7	53	CAC A His A	gtaagtaaagccaat…	intron 7	771	…ctcttcctcccctag	AT CTG sn Leu
Exon 8	32	AAC AAG Asn Lys	gtgagcactgccact…	intron 8	973	…ctctcatattttcag	GTG GAT Val Asp
Exon 9	1180						

15

Table 3 **Intron/exon lengths and similarities of *Ceacam1* and *Ceacam2***

Exon	Length (bp)		Similarity	Intron	Length (bp)		Similarity
	Ceacam1	Ceacam2			Ceacam1	Ceacam2	
1	304	304*	88.9%	1	767	766	97.4%
2	360	360	76.9%	2	1540	7807	29.7%
3	285	285	97.9%	3	467	465	94.7%
4	255	255	99.6%	4	1726	1717	97.9%
5	276	276	87.0%	5	5358	6737	59.8%
6	121	118	78.5%	6	796	798	98.5%
7	53	53	100%	7	587	771	64.5%
8	32	32	100%	8	948	973	94.3%
9	1180	1180+	86.8%				

*includes 240 bp of promoter and 5' untranslated sequence.

+includes 889 bp of 3' untranslated sequence.

14

FIGURE LEGENDS

Figure 1 Structure of the mouse *Ceacam1* and *Ceacam2* genes

Numbered open boxes indicate exons. The thick line connecting the exons represents introns. **(A)**

The maps of two overlapping λ clones (λ3 and λ7) that contain *Ceacam1*. **(B)** The maps of two

overlapping λ clones (λ13 and λ15) that contain *Ceacam2*. The restriction maps are aligned with

the intron/exon map. E, EcoRI; H, HindIII; X, XbaI. The enzyme sites in parentheses are from the

multiple cloning region of the λphage vector. **(C)** Comparison of the mouse *Ceacam1* and

Ceacam2 genes.

Figure 2 Comparison of nucleotide and protein sequences of *Ceacam1* and *Ceacam2* exon 2

Oligonucleotides for PCR analyses are underlined. Oligonucleotides for hybridization (Oligo 181

and Oligo 182) are boxed. The nucleotides are numbered with the A of the start ATG as the first

nucleotide according to the genomic sequence. The amino acids are numbered with the start

methionine (ATG) as the first amino acid.

Figure 3 Expression of *Ceacam1* and *Ceacam2* messages in various mouse tissues

(A) Specificity of Oligos 181 and 182 in distinguishing *Ceacam1* from *Ceacam2*. **(B and C)**

Tissue-specific expression of *Ceacam1* (B) and *Ceacam2* (C) analyzed by RT-PCR and

hybridization with Oligo 181 and Oligo 182, respectively. The positions of the PCR products are

indicated by the arrows on the right.

Figure 4 *Ceacam2*

17

(A) cDNA Structure. The numbered boxes indicate exons. The thick lines connecting the exons represent introns. The dashed lines show the portions of the exon sequences included in testis *Ceacam2* cDNA. **(B)** cDNA and protein sequence.

Figure 5   Expression of *Ceacam1* (A) and *Ceacam2* (B) messages during embryonic development

(A) Analysis of RT-PCR products of *Ceacam1* on an agarose gel and gene-specific hybridization using Oligo 181. **(B)** Analysis of RT-PCR products of *Ceacam2* on an agarose gel and gene-specific hybridization using Oligo 182. Marker sizes are indicated by thin arrows and the size of the PCR product is indicated by thick arrow.
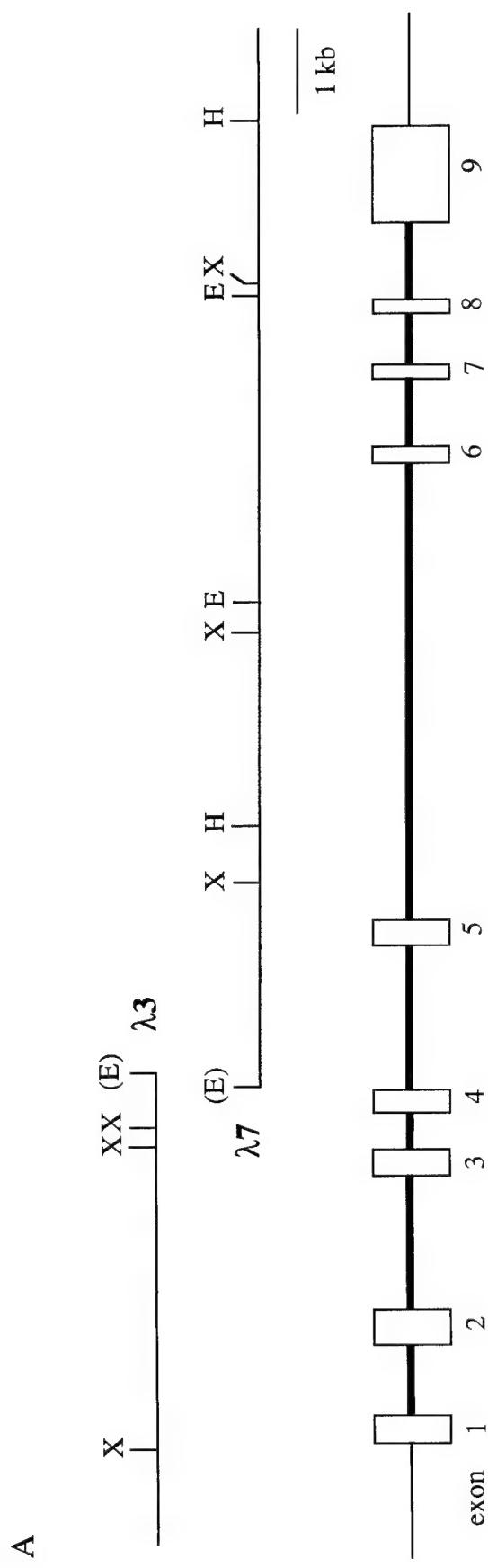
Figure 1

Figure 1

Figure 1

C
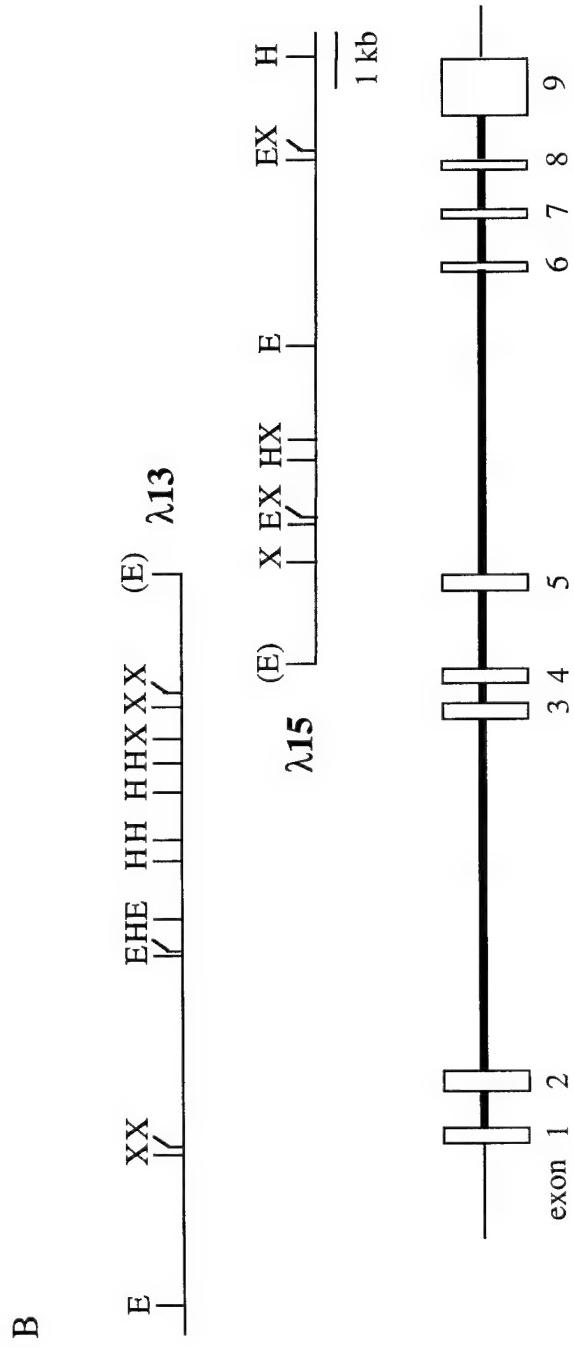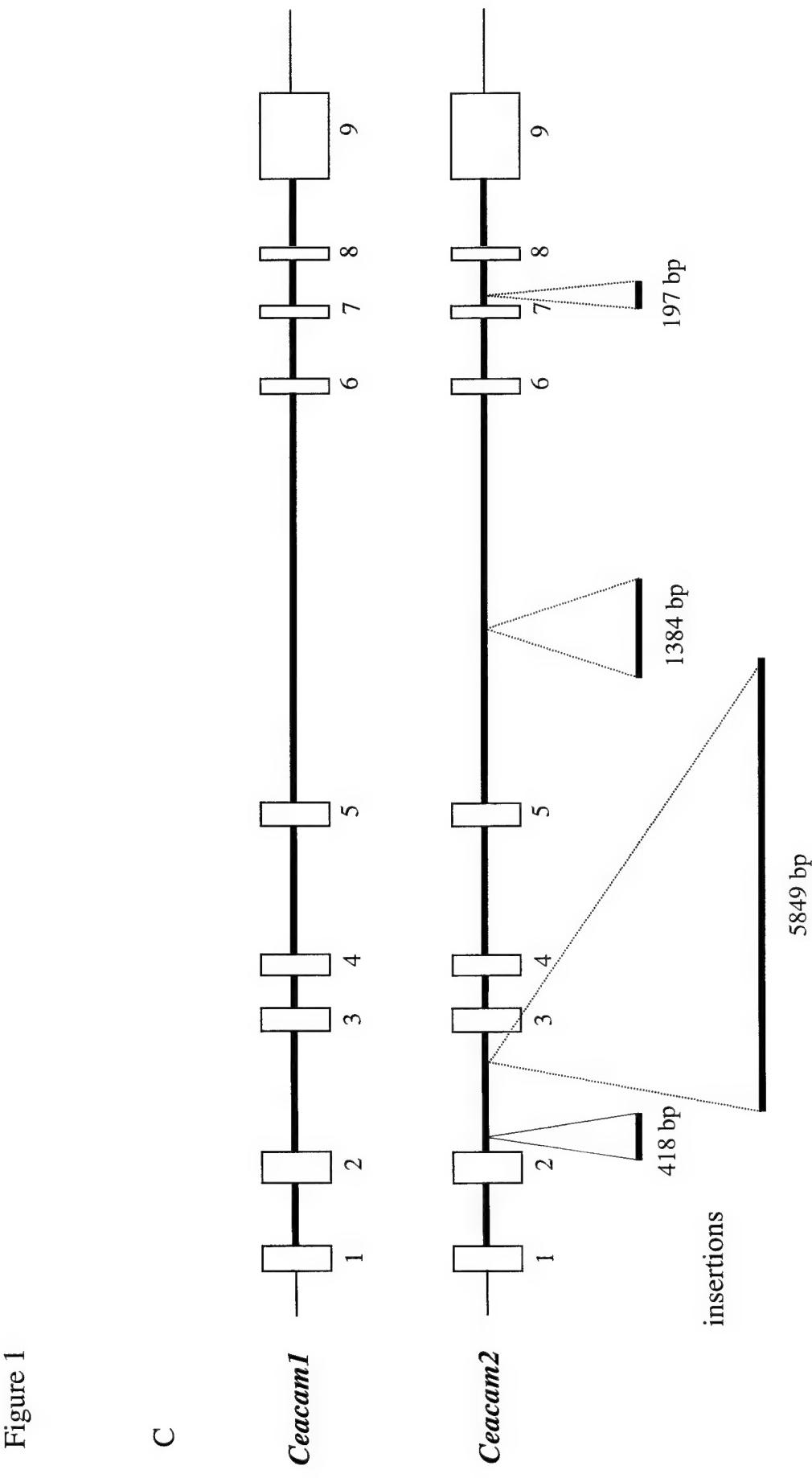
Figure 2

Exon 2 (nucleotide)

```
          843       853       863       873       883       893
Ceacam1 CCTCACTTTTAGCCTCCTGGAGCCCTGCCACCACTGCTGAAGTCACCATTGAGGCTGTGC
        :::::::::::::::::::::::::::: ::::::::::: :::: ::   ::  :::: : :
Ceacam2 CCTCACTTTTAGCCTCCTGGAGCCCTCCCACCACTGCACAAGTGACTGTTATGGCTTTTC
          840       850       860       870       880       890

          903       913       923       933       943       953
Ceacam1 CGCCCCAGGTTGCTGAAGACAACAATGTTCTTCTACTTGTTCACAATCTGCCCCTGGCGC
        : : ::: : ::::::::: :::::::::: ::::: ::::: ::::: ::       ::
Ceacam2 CACTCCACGCTGCTGAAGGCAACAATGTTATTCTAGTTGTTTACAATATGATGAAGGGAG
          900       910       920       930       940       950

          963       973       983       993       1003      1013
Ceacam1 TTGGAGCCTTTGCCTGGTACAAGGGAA|ACACTACGGCTATAGACAAAG|AAATTGCACGAT
        :    :::::::   :::: ::::::::::   :::::::: ::: : :    ::::::: :::::
Ceacam2 TCTTGGCCTTTAGCTGGCACAAGGGAT|CTACTACGTCTACAAATGCTG|AAATTGTACGAT
          960       970       980       990       1000      1010

          1023      1033      1043      1053      1063      1073
Ceacam1 TTGTACCAAATAGTAATATGAATTTCACGGGGCAAGCATACAGCGGCAGAGAGATAATAT
        ::::: ::    :  ::::: :: : : :    ::::  : : :::: :::::::::::  : :::
Ceacam2 TTGTAACAGGCACTAATAAGACTATAAAAGGGCCTGTACACAGTGGCAGAGAGGCACTAT
          1020      1030      1040      1050      1060      1070

          1083      1093      1103      1113      1123      1133
Ceacam1 ACAGCAATGGATCCCTGCTCTTCCAAATGATCACCATGAAGGATATGGGAGTCTACACAC
        :::::::::::::::::::::: ::::::: : :::::::::::::::::: ::::::::::::::::
Ceacam2 ACAGCAATGGATCCCTGCTCATCCAAAGGGTCACCATGAAGGATACGGGAGTCTACACAA
          1080      1090      1100      1110      1120      1130

          1143      1153      1163      1173      1183      1193
Ceacam1 TAGATATGACAGATGAAAACTATCGTCGTACTCAGGCGACTGTGCGATTTCATGTACACC
        :::: :::::::::: :: :::::::::::     :::::  : ::::::::::::::
Ceacam2 TAGAAATGACAGATCAAGACTATCGTCGTAGGGTCCTGACTGGACAATTTCATGTACACA
          1140      1150      1160      1170      1180      1190
```


Exon 2 (protein)

```
          32        42        52        62        72        82
Ceacam1 SLLASWSPATTAEVTIEAVPPQVAEDNNVLLLVHNLPLALGAFAWYKGNTTAIDKEIARF
        :::::::::: ::: ::  : :    ::  ::: : : :       ::  : :: ::    :: ::
Ceacam2 SLLASWSPPTTAQVTVMAFPLHAAEGNNVILVVYNMMKGVLAFSWHKGSTTSTNAEIVRF
          32        42        52        62        72        82

          92        102       112       122       132
Ceacam1 VPNSNMNFTGQAYSGREIIYSNGSLLFQMITMKDMGVYTLDMTDENYRRTQATVRFHVH
        :    :      :    :::: ::::::::: : ::::  :::: ::: ::: :  ::::
Ceacam2 VTGTNKTIKGPVHSGREALYSNGSLLIQRVTMKDTGVYTIXMTDQDYRRRVLTGQFHVH
          92        102       112       122       132
```

22

Figure 3

A

Oligo 181 (*Ceacam1*)

5.5 kb

Oligo 182 (*Ceacam2*)

6.6 kb

23

Fig. 3 (B)

β actin

prostate gland
uterus
ovary
adrenal gland
skin
testis
lung
muscle
small intestine
stomach
liver
thymus
spleen
kidney
heart
brain

1000
600
200

24

Fig. 3 (C)

β actin

brain
heart
kidney
spleen
thymus
liver
stomach
small intestine
muscle
lung
testis
skin
adrenal gland
ovary
uterus
prostate gland

1000
600
200

25

Figure 4

A

exon 1   2   3 4   5   6   7 8   9

Ceacam2 testis cDNA

1141 bp
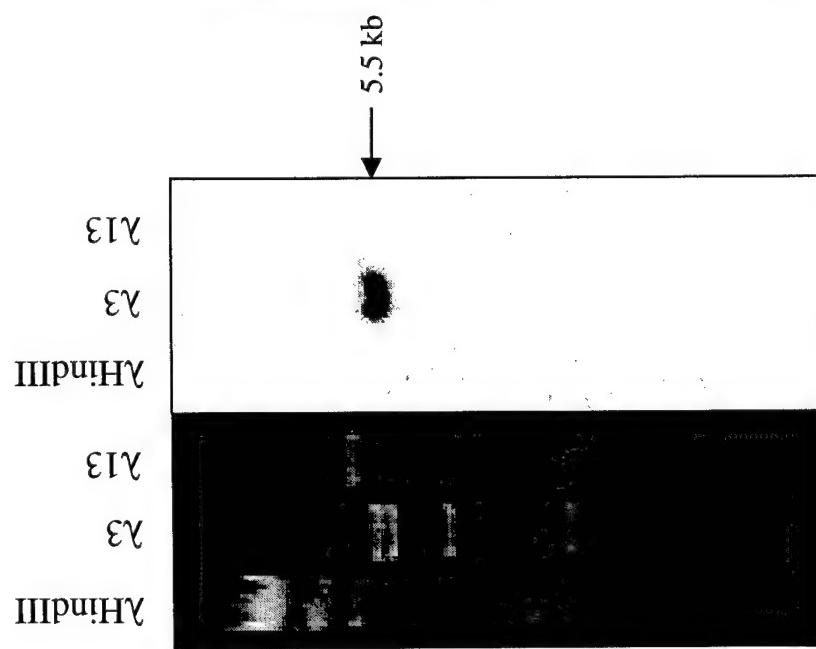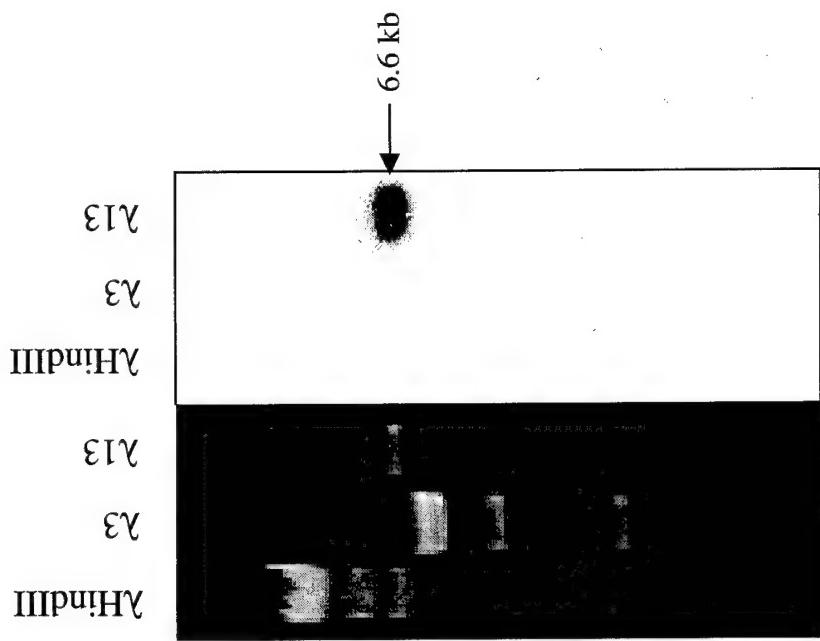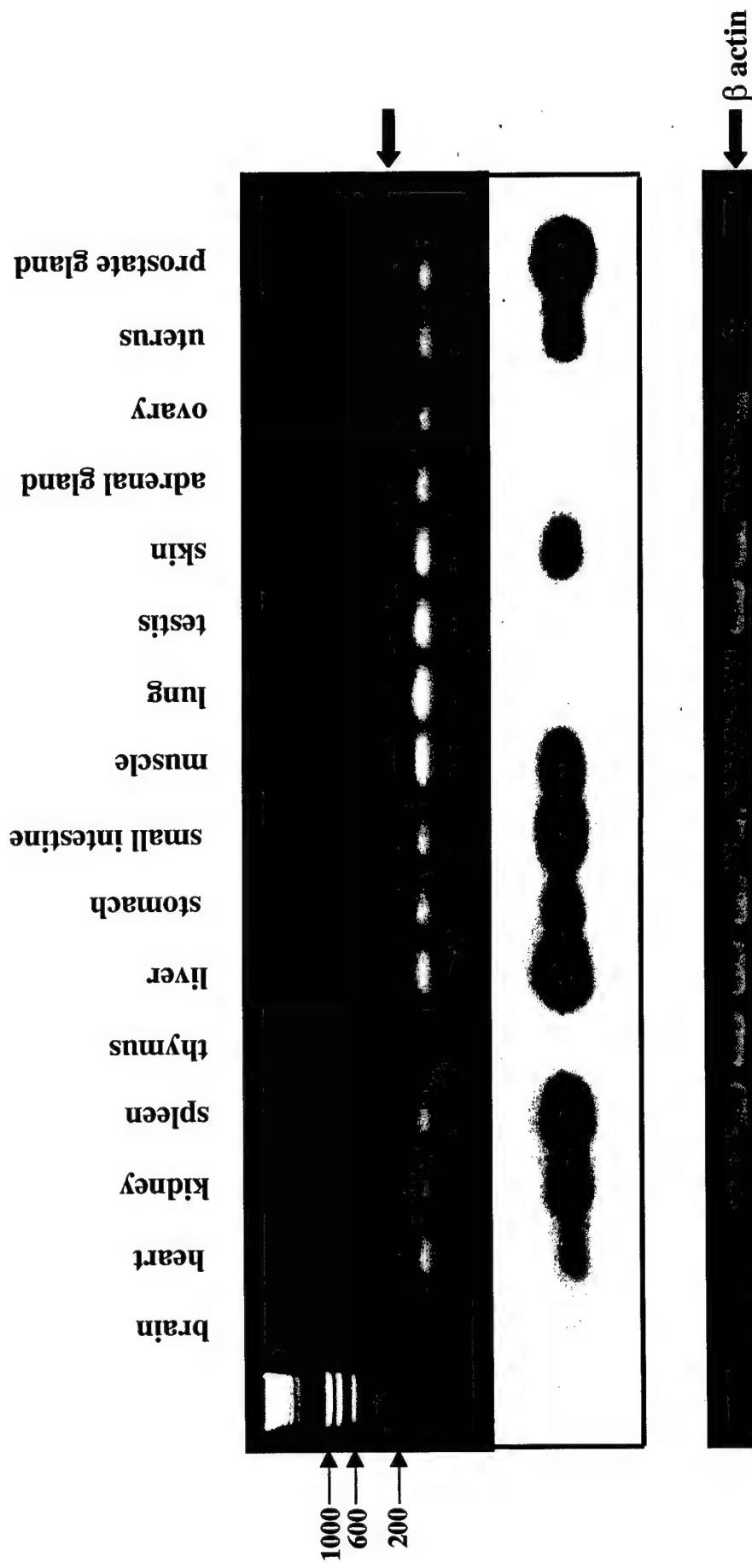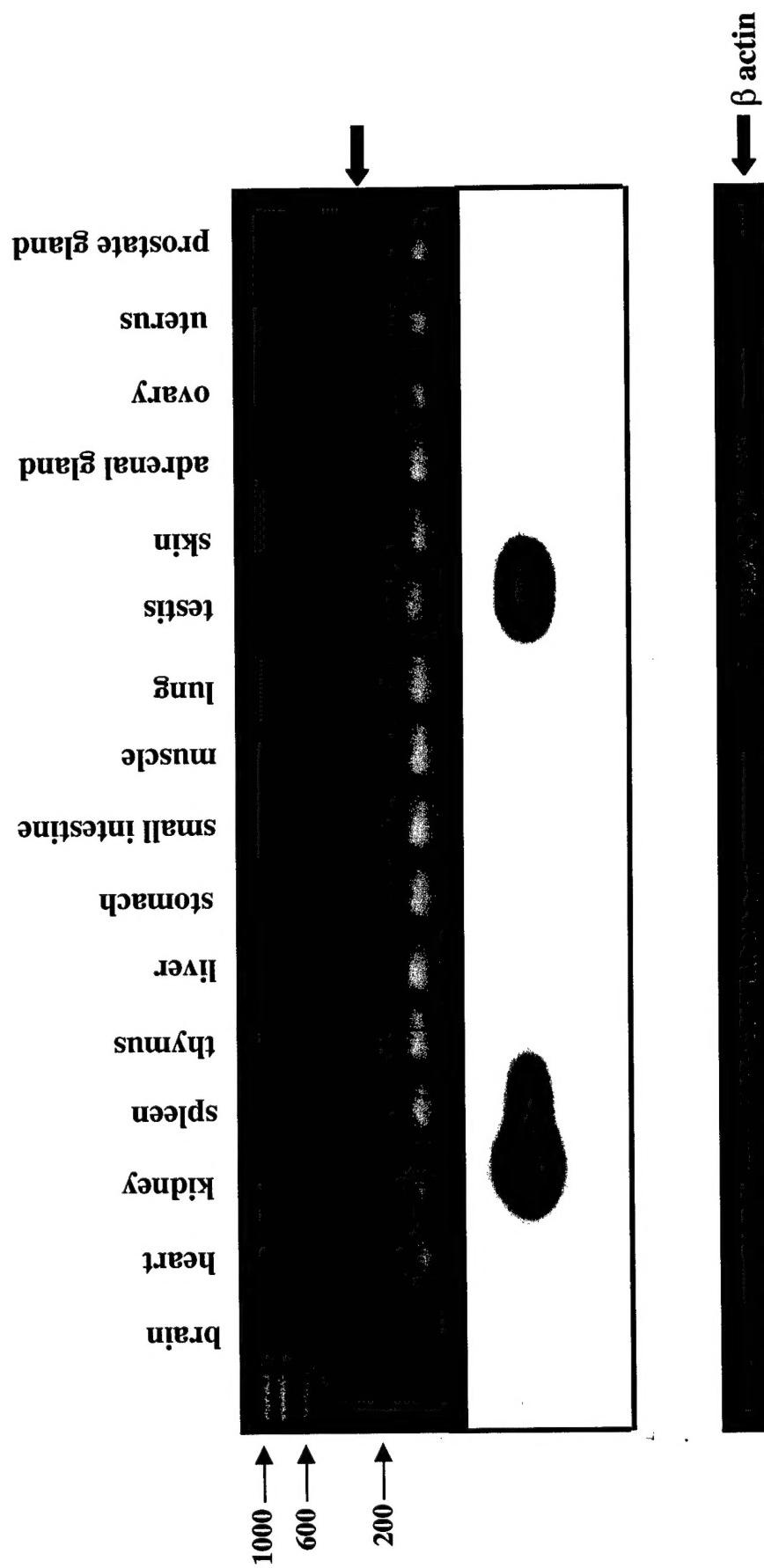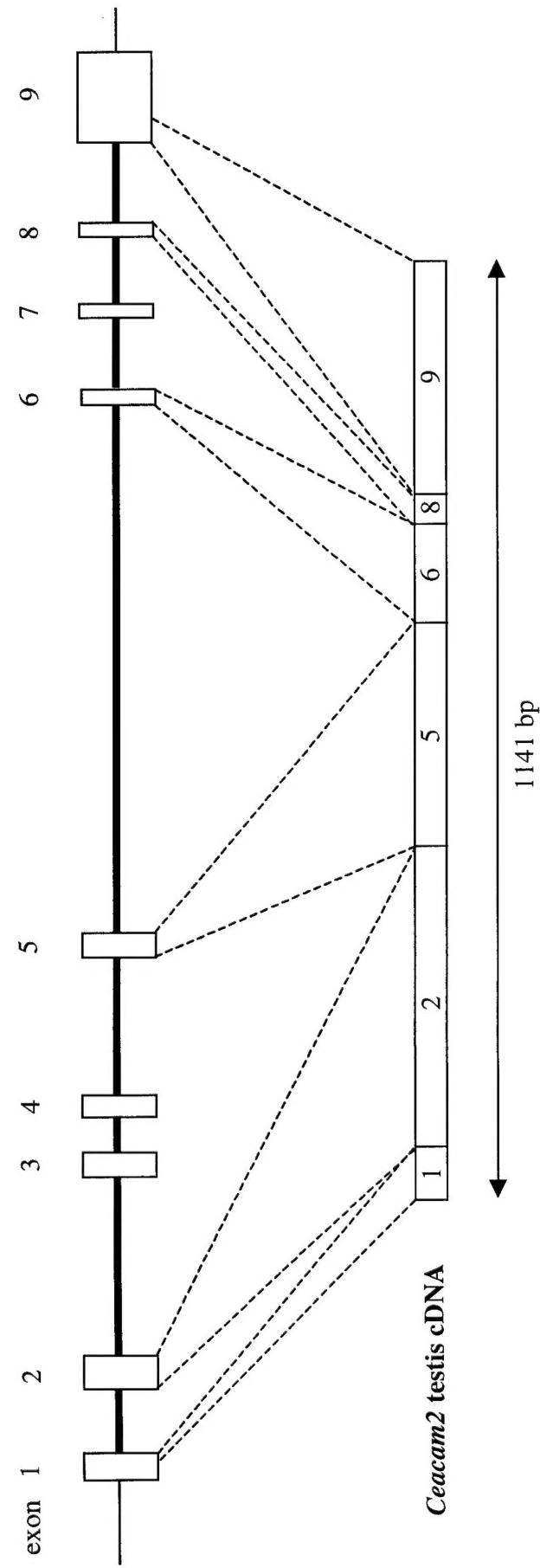
Figure 4

B

```
 -61  ATATCAGGGCAGCCAGGCTTAGCAGTAGTGTTGGAGAAGAAGCTAGCAGGCAGCAGAGAC   0

  1   ATGGAGCTGGCCTCAGCACATCTCCACAAAGGGCAGGTTCCCTGGTTTGGACTACTGCTC   60
      M  E  L  A  S  A  H  L  H  K  G  Q  V  P  W  F  G  L  L  L

 61   ACAGCCTCACTTTTAGCCTCCTGGAGCCCTCCCACCACTGCACAAGTGACTGTTATGGCT   120
      T  A  S  L  L  A  S  W  S  P  P  T  T  A  Q  V  T  V  M  A

121   TTTCCACTCCACGCCGCTGAAGGCAACAATGTTATTCTAGTTGTTTACAATATGATGAAG   180
      F  P  L  H  A  A  E  G  N  N  V  I  L  V  V  Y  N  M  M  K

181   GGAGTCTCGGCCTTTAGCTGGCACAAGGGATCTACTACGTCTACAAATGCTGAAATTGTA   240
      G  V  S  A  F  S  W  H  K  G  S  T  T  S  T  N  A  E  I  V

241   CGATTTGTAACAGGCACTAATAAGACTATAAAAGGGCCTGTACACAGTGGCAGAGAGACA   300
      R  F  V  T  G  T  N  K  T  I  K  G  P  V  H  S  G  R  E  T

301   CTATACAGCAATGGATCCCTGCTCATCCAAAGGGTCACCATGAAGGATACGGGAGTCTAC   360
      L  Y  S  N  G  S  L  L  I  Q  R  V  T  M  K  D  T  G  V  Y

361   ACAATAGAAATGACAGATCAAAGCTATCGTCGTAGGGTCCTGACTGGACAATTTCATGTA   420
      T  I  E  M  T  D  Q  S  Y  R  R  R  V  L  T  G  Q  F  H  V

421   CACAAGCCAGTGACTCAGCCCTCCCTCCAAGTCACCAACACCACAGTCAAAGAACTAGAC   480
      H  K  P  V  T  Q  P  S  L  Q  V  T  N  T  T  V  K  E  L  D

481   TCTGTCACCCTGACCTGCTTGTCGAAAGACCGTCAAGCCCACATCCATTGGATCTTCAAC   540
      S  V  T  L  T  C  L  S  K  D  R  Q  A  H  I  H  W  I  F  N

541   AATGATACTCTTCTAATCACAGAGAAGATGACAACCTCTCAGGCGGGACTCATCCTCAAA   600
      N  D  T  L  L  I  T  E  K  M  T  T  S  Q  A  G  L  I  L  K

601   ATAGACCCTATTAAGAGGGAAGATGCCGGCGAGTATCAGTGTGAAATCTCGAATCCAGTC   660
      I  D  P  I  K  R  E  D  A  G  E  Y  Q  C  E  I  S  N  P  V

661   AGCGTCAAGAGGAGCAACTCAATCAAACTGGAAGTAATATTTGACTCAACATATGACATT   720
      S  V  K  R  S  N  S  I  K  L  E  V  I  F  D  S  T  Y  D  I

721   TCAGATGTCCCCATTGCTGTAATCATAACTGGAGCTGTGGCCGGGGTGATTCTAATAGCA   780
      S  D  V  P  I  A  V  I  I  T  G  A  V  A  G  V  I  L  I  A

781   GGGCTGGCATATCGCCTCTGTTCCAGGAAGTCTCGCTGATCTGGCTCCTTCTGACAACTC   840
      G  L  A  Y  R  L  C  S  R  K  S  R  *

841   TCCTAACAAGGTGGATGACGTCGCATACACTGTCCTGAACTTCAATTCCCAGCAACCCAA   900

901   CCGGCCAACTTCAGCCCCTTCTTCTCCAAGAGCCACAGAAACAGTTTATTCAGAAGTAAA   960

961   AAAGAAG
```

27

Fig. 5

(A) Ceacam1

(B) Ceacam2

embryo/8.5 d
embryo/9.5 d
embryo/12.5 d
embryo/19 d

600 →
200 →
actin →

28